CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER for: 020779, S022

STATISTICAL REVIEW(S)

approval for marketing on March 14, 1997 for the treatment of HIV infection. The approval was based on results of changes in surrogate markers (CD_4 count and plasma HIV RNA) in patients for up to 24 weeks of therapy from 2 well controlled Phase III clinical trials: AG1343-506 and AG1343-511. The present NDA supplement contains 48 week follow-up data from these 2 trials and 48 week results from a trial evaluating twice-a-day dose administration (AG1343-542). The sponsor seeks to obtain full approval status for the twice-a-day dosing regimen.

There is no new nonclinical pharmacology/toxicology information submitted with the present 2 NDA's. No changes are made in the "Carcinogenesis and Mutagenesis," "Pregnancy, Fertility, and Reproduction," and "Nursing Mothers" sections of the Label.

CONCLUSION

Since no change was made in the "Carcinogenesis and Mutagenesis," "Pregnancy, Fertility, and Reproduction," and "Nursing Mothers" sections of the Label and no new nonclinical pharmacology/toxicology information was submitted, there is no regulatory action associated with this review from the nonclinical pharmacology/toxicology perspective.

/S/

Ita Yuen, Ph.D.
Reviewing Pharmacologist

CONCURRENCE

HFD-530/WDempsey /S/ 11/26/99
HFD-530/JFarrell /S/ 11/24/99

CC

HFD-530/NDA 20,778
HFD-530/NDA 20,779 (022)
HFD-530/Division File
HFD-530/SLynch
HFD-345
HFD-530/TWu
HFD-530/LIaconno-Connors
HFD-530/GLunn
HFD-530/RKumi
HFD-530/THammerstrom

STATISTICAL REVIEW AND EVALUATION

SNDA#:	20-778/S-011, 20-779/S-022
APPLICANT:	Agouron Pharmaceuticals, Inc
NAME OF DRUG:	Viracept®
INDICATION:	Treatment of HIV Infection
TYPE OF REVIEW:	Clinical
DOCUMENTS REVIEWED:	Volumes 18.61, 20.61, 21.61
MEDICAL INPUT:	Teresa Wu, M.D. (HFD-530)

STATISTICAL REVIEW AND EVALUATION

NDA#:

20-779

APPEARS THIS WAY
ON ORIGINAL

1. Background

1.1 Objectives in Trial

The applicant submitted one randomized, open-label, controlled clinical trials with Viracept (nelfinavir, NFV) for this supplement, trial 542.

The primary objective of this study was to compare the efficacy of tid NFV therapy at a dose of 750 mg to that of bid NFV therapy at doses of 750 mg, 1000 mg, or 1250 mg when used in combination with 3TC at 150 mg bid and d4T at 40 mg bid. During the trial, the 750 mg and 1000 mg bid groups had their doses increased to 1250 mg and the trial was modified to a comparison of the two remaining arms: 750 mg tid and 1250 mg bid.

The primary efficacy endpoint in the study was percent below limit of quantitation (LOQ) at 48 weeks. The study population was HIV-1 infected patients with less than 6 months of anti-retroviral therapy and with HIV RNA levels > 15 k copies/ml.

1.2 Summary of Study Designs

Study 542 was open-label, randomized, two-arm (initially, four arm), parallel, active controlled, multi-center trials conducted in Europe. (The trial was not blinded because of the difference in dosing schedules. During the four-arm phase, subjects assigned to one of the bid arms were blinded as to which dose they were assigned.)

Initially, subjects were randomly assigned in a 1:1:1:1 ratio to one of the three bid arms or to the tid arm. Enrollment began on March 11, 1997. Trial 511 subsequently showed that the 750 mg tid dose (= 2250 mg daily) was superior to the 500 mg tid dose = (1500 mg daily). When these results became known on Sept. 18, 1997, subjects on the two lower bid doses, with daily doses $\leq$ 2000, were re-assigned to the 1250 mg bid arm. New subjects enrolling after this date were enrolled in a 1:1 ratio to the 1250 mg bid or the 750 mg tid arm. Randomization was stratified by CD4 level ($\leq$ 200 and >200 cells/mm^3) and by center.

The applicant introduced the term treatment failure to mean an increase of HIV RNA to greater than three times the patient's nadir and as increase of HIV RNA to greater than 10K copies/mL, confirmed by a second visit. It is unclear whether HIV RNA had to increase to exceed both thresholds or either threshold. In the event of treatment failure, the patient was to begin a new therapy. It is important to note that this definition of treatment failure was not used in the statistical analysis of the data. Time to failure and percent failed at week 48, according to this definition of failure, were neither primary nor secondary statistical endpoints. Because this is not an endpoint, this event should not be called treatment failure and will not be called treatment failure in this review. In the remainder of this review, the term treatment failure will refer to the earliest of viral rebound to > LOQ or treatment discontinuation or patient loss to follow-up.

The decision to discontinue the lower dose arms was based on data from an external source (trial 511) rather than from an interim analysis in this trial.

1.3 Patient Accounting and Baseline Characteristics

459 patients were enrolled in the trial 542, beginning on Mar 11, 1997, and continuing until 48 weeks before Nov 5, 1998. Of these, 455 patients were considered evaluable for safety and 447 were considered evaluable for efficacy. Of these 447 patients, 72 discontinued prior to the preparation of the preliminary report. The subjects were enrolled at 21 centers in Europe.

In trial 542, the study population was 84% male with a mean age of 36 years. They were 90% white and 7% black. The mean duration of HIV was 36 months; the mean CD4 count at baseline was 280 cells/mm^3; the mean HIV RNA level was 5.0 logs.

Table 1.3 A summarizes the patient status at the time of the preliminary analysis as reported by the applicant, including the primary reasons for discontinuation from treatment. There are a

above LOQ or if they had previously discontinued their assigned therapy. Subjects with missing observations at the target week were considered as missing if both preceding and succeeding observations were BLQ or if they had not been enrolled long enough to reach the target week at the time of the preliminary analysis. Time until a confirmed return of HIV RNA to > LOQ was a secondary endpoint.

1.5 Summary of Statistical Analysis

The percent BLQ at each week was computed, using the rules in section 1.4.2 for dealing with missing values. A simple normal theory based confidence interval for the difference between TID and pooled BID arms was computed, ignoring all stratifying variables used in the randomization. The applicant considered a 95% confidence that the BID arm was no more than 12.5% worse than the TID arm to be an establishment of clinical equivalence.

2. Summary of Applicant's Results

The results for the comparisons in trial 542 are given in table 2.1 A.

TABLE 2.1 A
HIV RNA RESULTS IN TRIAL 542
PERCENT BLQ BY ARM AND
CONFIDENCE LIMITS FOR DIFFERENCE BETWEEN ARMS

Endpoint	% BLQ BID	% BLQ TID	95% C.I. for TID % - BID %
Week 24	180/240 = 75%	84/108 = 78%	(-12.3%, 6.8%)
Week 48	131/204 = 64%	48/75 = 64%	(-12.5%, 12.9%)

3. Summary of Applicant's Conclusions

The applicant concluded that this trial showed that improvements in HIV RNA levels were similar in magnitude and durability over time between 750 mg TID and 1250 mg BID Viracept when used in conjunction with other anti-retroviral drugs.

4. Statistical Reviewer's Comments and Analyses

The applicant's analysis of trial 542 suffers from several defects. First, their analysis of percent BLQ is based on incorrect algorithm which fails to adjust for the occurrence of occasional false negative and false positive HIV RNA assay results. Second, their analyses of percent BLQ are inappropriate because they are not stratified by the variables used to stratify the randomization. Third, because this is an open label trial, one needs to determine the robustness of the conclusions to the effects of possibly biased treatment discontinuations.

Section 4.1 below will report the disposition of patients in both, including details for each dose in the BID arms. Section 4.2 will present the analysis of percent BLQ in both trials, using corrected times to failure for each patient, and using appropriate stratifications. Furthermore, this section will show that the observed p-values are small enough to conclude statistical significance.

Section 4.2 will also examine the results of analyses which test whether the conclusions are sensitive to assumptions that there was no differential behavior between subjects on the BID and TID arms and show that the amount of uncertainty about the difference between the two schedules increases when one takes into account the possibility of bias. If one takes reasonable precautions against biased results, one cannot be confident that the BID arm is not as much as 14% worse than the TID arm.

Section 4.3 will show that analysis of the secondary endpoint of time to loss of BLQ leads to the same conclusion of statistical significance. Thus, this conclusion is not sensitive to choice of endpoint.

4.1 Patient Dispositions

Tables 4.1 A and B give the patient dispositions in trial 542, as computed from the applicant's computer files. These tables show the results for the BID arms broken down according to the dose originally assigned and, within the 1250 mg arm, by before or after the decision to increase the 1000 and 750 mg doses to 1250 mg. Table 4.1 A gives reason for stopping treatment. Here, censoring refers specifically to administrative censoring: subjects were still on treatment but had not reached the target endpoint prior to 8 weeks before July 31, 1998. (Since visits were scheduled only once every 8 weeks, the FDA reviewer did not count a visit as missed if it was scheduled within 8 weeks of the date the database was closed for the preliminary analysis.)

TABLE 4.1 A
REASON FOR STOPPING TREATMENT
TRIAL 542

REASON	TID 750	BID 1250 LATE	BID 1250 EARLY	BID 1000	BID 750	ALL BID
Randomized	167	93	71	71	69	304
Completed 48 wks	57	0	54	53	53	160
Censored						
Before Week 24	54	53	0	0	0	53
Between Week 24 and 48	29	27	0	0	1	28
Toxicity	3	3	4	2	2	11
Intercurrent Illness	1	3	0	3	3	9
Other	23	7	13	13	10	43

Table 4.1 B gives the times of last observations of HIV RNA, regardless of whether subjects had discontinued their assigned treatments or not.

TABLE 4.1 B
TIME OF DISCONTINUATION

TRIAL 542

STOPPED	TID 750	BID 1250 LATE	BID 1250 EARLY	BID 1000	BID 750	ALL BID
Randomized	167	93	71	71	69	304
Before Week 24	71	63	7	6	5	81
Before Week 12	39	33	3	5	3	44
Between Week 12 and 24	32	30	4	1	2	37
Between Week 24 and 48	37	30	9	10	9	58
After Week 48	59	0	55	55	55	165

4.2 Analyses of Percent Below Limit of Quantitation:

The primary efficacy endpoint in this trial is percent BLQ at weeks 24 and 48. The former time point is used for accelerated approval, the latter for traditional approval. There are four problems with the data that preclude one from estimating the percent BLQ simply by counting the number of subjects above and below limit of quantitation at the target week. These four problems are as follows. 1) With preliminary reports on trials still continuing, some subjects still BLQ have not reached the target week. 2) Among those subjects enrolled more than 24 or 48 weeks before the preliminary closure of the database, not all are followed on assigned therapy to the target week. 3) Some subjects miss scheduled visits even though they have not been lost to follow-up. 4) The assays exhibit occasional false negative and false positive readings.

The correct estimate of percent BLQ is obtained from the Kaplan-Meier curve for time to failure, when failure is defined as the earlier of permanent rebound of viral load to levels above LOQ and switch from assigned therapy. The Kaplan-Meier estimates of percent not yet failed at weeks 24 and 48 are the primary endpoints. The corresponding 95% confidence limits for the differences between the active and control curves are used to conduct the statistical tests.

The four problems just mentioned require specific rules for dealing with missing data and with occasional high or low assay values. The FDA reviewer used the following rules in the analysis.

First, missing values were discarded from the analysis. Then time to virologic failure was defined as the first of two consecutive HIV RNA measurements above LOQ after a measurement at week 4 or later that was BLQ. This rule was applied even if one or more scheduled visits between the two values above LOQ were missing. Subjects with a single HIV RNA measurement above LOQ at their last visit were considered to have failed at their last visit regardless of why no later visits were made.

Second, time until virologic failure was defined as zero for subjects who were never BLQ, as time of the progression for subjects who experienced clinical progression while still BLQ, as time of the switch for subjects who switched their assigned therapy while still BLQ, and as time of the loss to follow-up for subjects who were lost to follow-up while still BLQ. Loss to follow-up means termination of HIV RNA measurements even if other observations on the patient continued.

Finally, failure times are considered censored for subjects who were still BLQ and on their assigned therapy and making scheduled visits for HIV RNA assays at the time that the database was closed for the preliminary analysis. All other times are considered to be observed failures.

There are other possible algorithms for dealing with missed HIV RNA assays, such as imputing all missed scheduled assays to be above LOQ. Discussion among FDA clinical and statistical reviewers produced a consensus that the algorithm adopted above is preferable to other imputations. In particular, a single HIV RNA assay above LOQ, followed by a missing scheduled assay, followed by further assays below LOQ would have been regarded as a failure by an algorithm that imputed missing values as above LOQ but was not regarded as a failure by the analyses used in this review.

Analyses based on simply counting number of subjects BLQ, number above LOQ, and taking ratios are unacceptable because there are occasional instances of subjects with a single assay above LOQ in the middle of a sequence of BLQ readings or with a single BLQ assay after a confirmed viral rebound. Treating these subjects as failures or successes if the single anomalous assay happens to occur at a target week produces manifestly incorrect answers. Furthermore, analyses based on counting rather than the Kaplan-Meier curve do not adequately handle missing data and administratively censored subjects.

There are three additional issues to be addressed in the estimation of percent BLQ. Of these three, the one with the most serious potential consequences is that this trial is open-label. One must thus be concerned about the effects of differential switching while still BLQ or differential loss to follow-up. In this trial, the FDA reviewer found only four TID subjects who were classed as failures because of discontinuations prior to viral rebound on the Amplicor assay. Using the Ultrasensitive assay, only one TID subject was called as a failure because of discontinuation before viral rebound. Thus, biased discontinuations do not appear to have made the BID arm look significantly better. The results reported below all use the conservative assumption that this TID subjects with pre-mature discontinuation was counted as censored rather than failed at time of discontinuation while the BID subjects with pre-mature discontinuations were counted as failures.

The second issue is also a consequence of the censoring occurring from the preliminary nature of these data. 27 TID and 35 BID subjects were administratively censored before reaching week 12 on treatment and without ever reaching BLQ on the Amplicor assay. Such subjects are counted as failures at time zero but they might not have had a fair chance to get below LOQ. To examine the sensitivity of conclusions to such artificially early failures, the FDA reviewer conducted two analyses in which these subjects were 1) counted as failures at time zero and 2) discarded from the analysis.

The third issue is that the applicant used a randomization

stratified by 2 levels of baseline CD4 count crossed with site. The analyses should reflect these stratifications. Consequently, the FDA reviewer computed separate Kaplan-Meier estimates for each of the baseline CD4 strata and then computed a pooled estimate, using a Mantel-Haenszel weighting. (In order to keep the pooled Kaplan-Meier curve decreasing, the same weights were used at each time point. These weights were the initial sample sizes in each stratum.)

The correct analysis should stratify on both site and baseline CD4 count. The FDA reviewer has computed the Mantel-Haenszel weighted average of percent BLQ at weeks 24 and 48, using separate Kaplan-Meier curves for each of the 40 strata defined by two levels of baseline CD4 count and site. This analysis results in the loss of some data because strata which do not have subjects on both TID and BID contribute nothing to the estimates. The results from this analysis are also included in the following tables.

It is also desirable that final summary statistics should display percent BLQ as a ratio of subjects observed BLQ to total subjects observed. The Kaplan-Meier curve does not provide such a ratio. If there are subjects who are censored before the target week, one could a) count those subjects as failed, b) count those subjects as BLQ, or c) discard those subjects from numerator and denominator in computing percent BLQ as a ratio. The Kaplan-Meier does none of these. Instead it counts a fraction of the censored subjects as failed. Specifically, it assumes that the censored subjects have the same failure rate as was actually observed in the non-censored subjects. Thus, if one tried to display a Kaplan-Meier percentage as a ratio of counts, the numerator and denominator would be fractions.

In order to provide an estimate based on a ratio of counts, the FDA reviewer also used method c) from the previous paragraph to compute percent BLQ as a ratio of counts. A Mantel-Haenszel test stratified by baseline CD4 level (≤ 200 or > 200) was used to compare the percent of subjects BLQ at week 24 on the TID arm to the pooled BID arm.

The results for both Kaplan-Meier estimates and direct ratios for [redacted] assays are given in table 4.2 A. This table gives the count and percent of BLQ subjects in each stratum and treatment arm. Table 4.2 B gives the 95% confidence intervals for the difference in percent BLQ between each of the Viracept arms and the placebo arm. The intervals are also computed using a Mantel-Haenszel pooling of the estimates within each stratum, not from the pooled counts.

TABLE 4.2 A
PERCENT BLQ AT WEEK 24, TRIAL 542
STRATIFIED BY BASELINE CD4 LEVEL

	BID	TID
Randomized	304	167
Kaplan-Meier Estimate		
Percent BLQ on [redacted]		
Censor<Wk 12 Included	69%	69%
Censor<Wk 12 = Missing	77%	83%
Stratified by Site & CD4	78%	83%
Percent BLQ on Ultrasensitive		
Censor<Wk 12 Included	60%	52%
Censor<Wk 12 = Missing	69%	65%
Stratified by Site & CD4	70%	66%
Using Counts		
Percent BLQ on [redacted]		
Censor<Wk 12 Included		
CD4 ≤ 200	60/105 = 57%	30/53 = 57%
CD4 > 200	122/168 = 73%	56/83 = 67%
Pooled	182/273 = 67%	86/136 = 63%
Censor<Wk 12 = Missing	182/238 = 76%	86/109 = 79%
Percent BLQ on Ultrasensitive		
Censor<Wk 12 Included		
CD4 ≤ 200	55/112 = 49%	25/60 = 42%
CD4 > 200	112/176 = 63%	50/96 = 52%
Pooled	167/288 = 58%	75/156 = 48%
Censor<Wk 12 = Missing	167/246 = 68%	75/120 = 63%

One can notice an apparent anomaly in the percents BLQ as computed directly from the counts: the denominators are larger for the Ultrasensitive assay than for the [redacted] assay. These subjects were all below 400 on [redacted] but were short of week 24 when the database was closed for the preliminary report. They are thus censored, which means that they are not included in the analysis in the estimates based on counts. These same subjects also remained above 50 on the Ultrasensitive assay and are thus counted as failures at time zero in the analyses based on that assay. In the Kaplan-Meier estimates, these censored subjects are included in the estimation on both assays, which is another reason why the Kaplan-Meier is the correct method for estimating percents BLQ. Most importantly for this trial, one can see that these small anomalies have little influence on the overall conclusions.

TABLE 4.2 B
PERCENT BLQ AT WEEK 24, TRIAL 542
95% INTERVALS FOR BID - TID DIFFERENCE
STRATIFIED BY BASELINE CD4 LEVEL

	BID - TID
Kaplan-Meier Estimate	
Percent BLQ on [redacted]	
Censor<Wk 12 Included	(-9%, 8%)
Censor<Wk 12 = Missing	(-14%, 3%)
Stratified by Site & CD4	(-13%, 2%)
Percent BLQ on Ultra	
Censor<Wk 12 Included	(-1%, 18%)
Censor<Wk 12 = Missing	(-6%, 14%)
Stratified by Site & CD4	(-5%, 13%)
Using Counts	
Percent BLQ on [redacted]	
Censor<Wk 12 Included	(-6%, 13%)
Censor<Wk 12 = Missing	(-12%, 7%)
Percent BLQ on Ultra	
Censor<Wk 12 Included	(-0%, 20%)
Censor<Wk 12 = Missing	(-5%, 15%)

One can see from these tables that the percent BLQ on the Ultrasensitive assay is approximately the same on the BID and the

TID arms and with 95% confidence, the BID arm is no more than 6% worse than the TID arm. The results with the [redacted] assay are similar with one noteworthy difference. If subjects who are administratively censored early are counted as censored rather than as failures, then there is more statistical uncertainty about the difference between the TID and BID arms. With 95% confidence, one cannot rule out the possibility of the BID arm being as much as 14% worse than the TID arm. The FDA reviewer notes that this ambiguity could be eliminated by simply waiting to collect longer term data on the 46 subjects who were censored without a reasonable chance to achieve viral loads below LOQ.

One can also see that the size of the treatment effect is consistent across baseline CD4 strata for both assays. Both BID and TID doses had slightly worse outcomes in the stratum with poorer baseline CD4 levels but there is no evidence of a treatment-baseline CD4 interaction.

Because this report is preliminary, there is extensive administrative censoring by week 48. The results for week 48 are summarized in tables 4.2 C and D. Point estimates are given in table 4.2 C and 95% confidence intervals for the differences between the BID and TID arms are given in table 4.2 D.

TABLE 4.2 C
PERCENT BLQ AT WEEK 48, TRIAL 542
STRATIFIED BY BASELINE CD4 LEVEL

	BID	TID
Randomized	304	167
Kaplan-Meier Estimate		
Percent BLQ on [ ]		
Censor<Wk 12 Included	57%	60%
Censor<Wk 12 = Missing	64%	72%
Stratified by Site & CD4	66%	75%
Percent BLQ on Ultrasensitive		
Censor<Wk 12 Included	44%	38%
Censor<Wk 12 = Missing	51%	48%
Stratified by Site & CD4	52%	52%
Using Counts		
Percent BLQ on [ ]		
Censor<Wk 12 Included	129/247 = 52%	48/106 = 45%
Censor<Wk 12 = Missing	129/212 = 61%	48/79 = 61%
Percent BLQ on Ultrasensitive		
Censor<Wk 12 Included	105/266 = 39%	37/133 = 28%
Censor<Wk 12 = Missing	105/224 = 47%	37/97 = 38%

TABLE 4.2 D
PERCENT BLQ AT WEEK 48, TRIAL 542
95% INTERVALS FOR BID - TID DIFFERENCE
STRATIFIED BY BASELINE CD4 LEVEL

	BID - TID
Kaplan-Meier Estimate	
Percent BLQ on ( )	
Censor<Wk 12 Included	(-14%, 7%)
Censor<Wk 12 = Missing	(-19%, 3%)
Stratified by Site & CD4	(-18%, 1%)
Percent BLQ on (Ultra)	
Censor<Wk 12 Included	(-4%, 16%)
Censor<Wk 12 = Missing	(-9%, 15%)
Stratified by Site & CD4	(-10%, 10%)
Using Counts	
Percent BLQ on ( )	
Censor<Wk 12 Included	(-5%, 18%)
Censor<Wk 12 = Missing	(-13%, 12%)
Percent BLQ on (Ultra)	
Censor<Wk 12 Included	(2%, 21%)
Censor<Wk 12 = Missing	(-3%, 20%)

The percent of BLQ in both Viracept arms decreased slightly at week 48, relative to the values at week 24. The pattern of differences between BID and TID, however, remained the same. There was an increase in the amount of uncertainty about the differences. With 95% confidence, the BID was no worse than 4-9% worse than the TID arm on percent BLQ with the (Ultrasensitive) assay. With respect to percent BLQ on the ( ) assay, the BID arm might with 95% confidence be as much as 14-19% worse than the TID arm.

4.3 Analyses with Time to Virologic Failure

The FDA reviewer also computed Kaplan-Meier estimates for time until virologic failure comparing the pooled BID and TID arms. The FDA reviewer also computed log-rank tests and Kaplan-Meier graphs comparing the three BID doses to each other.

The Kaplan-Meier curves for times to virologic failure is given in figure 4.3 i. This figure gives the results for the [redacted] assay. Because this is an open label study, the reviewer used the results in which subjects lost to follow-up are failures on the BID arm and censored on the TID arm. Results for the Ultrasensitive assay looks similar. One can see that, even after adjusting for potential biases due to lack of blind, the two dose regimens look similar.

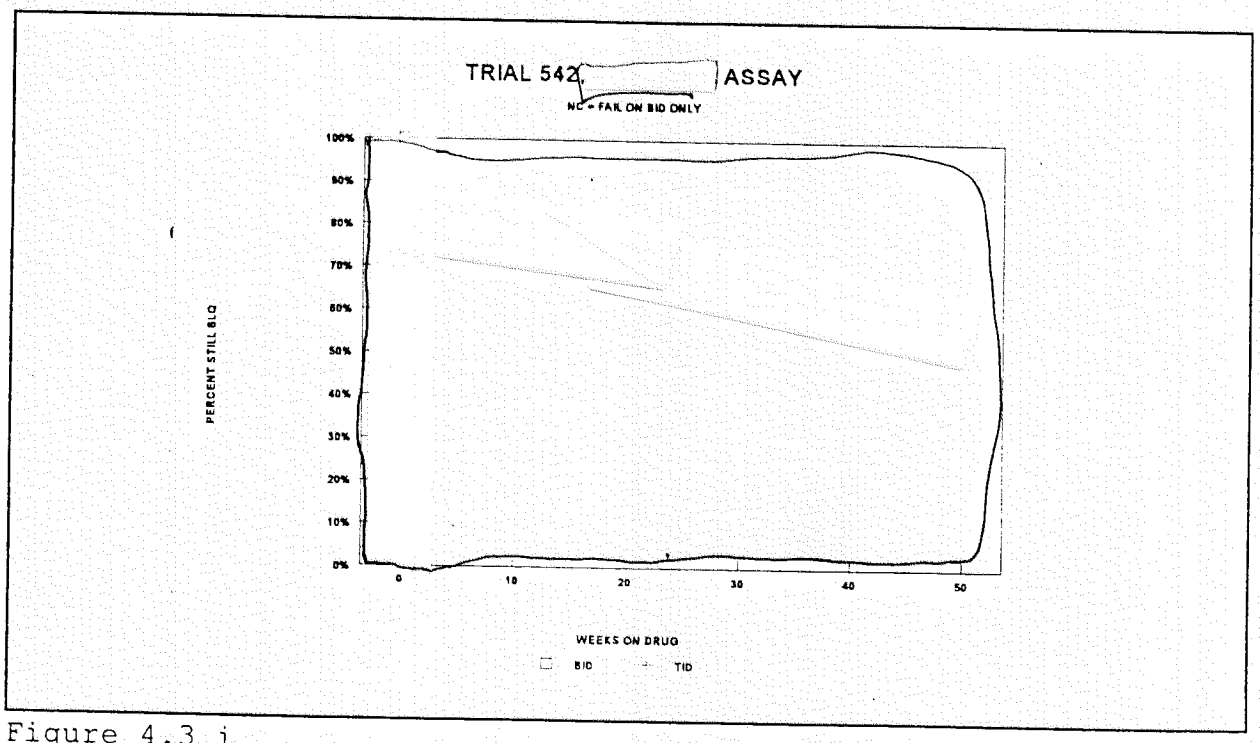

Figure 4.3 i

Confidence intervals for the differences between the TID and BID arms are shown in figure 4.3 ii for the [redacted] assay. The most important feature of this plot is that beyond week 30, one cannot be 95% confident that the BID arm is not 15% worse in percent BLQ than the TID arm. The lower confidence bounds look worse when one also includes the uncertainty due to counting subjects as failures when they are censored so early that they don't have time to reach BLQ. The lower confidence bounds look better when the Ultrasensitive assay is used as the standard.

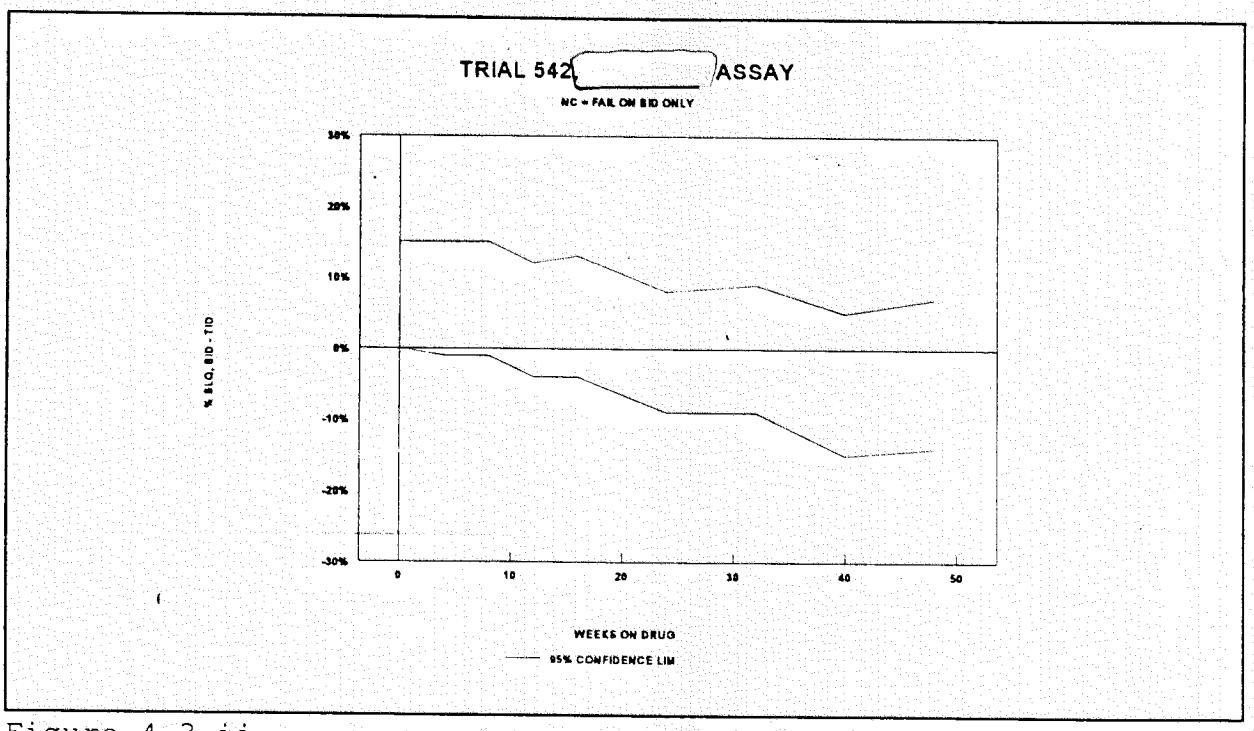

Figure 4.3 ii

As discussed in section 4.2, these Kaplan-Meier curves are actually obtained as Mantel-Haenszel weighted averages of separate Kaplan-Meier curves computed in each baseline CD4 stratum. The curves for each stratum (not reproduced here) showed only slight differences among the strata, a result compatible with the results given in table 4.2 B above using direct counts. As might be expected, there was a suggestion of higher percents BLQ with higher baseline CD4 counts.

5. Statistical Reviewer's Summary

The results from trial 542 show that, with 95% confidence, the BID regimen was no more than 14% worse than the TID regimen in percent BLQ at week 24 and no more than 19% worse at week 48. This uncertainty includes adjustment for the potentially biased results of open label assignment and for the uncertainty about the responses of subjects censored too early to accurately determine their failure times. The size of the uncertainty in